

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 June 2001 (14.06.2001)

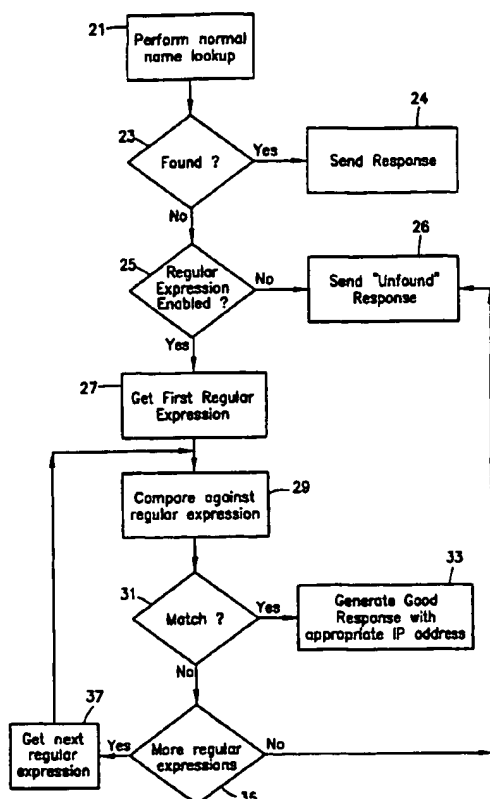
PCT

(10) International Publication Number
WO 01/42943 A1

- (51) International Patent Classification⁷: **G06F 15/16** (74) Agents: **McCABE, Philip, J. et al.**; Kenyon & Kenyon, 333 W. San Carlos Street, Suite 600, San Jose, CA 95110 (US).
- (21) International Application Number: **PCT/US00/09095**
- (22) International Filing Date: **5 April 2000 (05.04.2000)** (81) Designated States (*national*): AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW.
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:
09/457,420 **7 December 1999 (07.12.1999)** **US**
- (71) Applicant: **ABRA INCORPORATED [US/US]**; 4 North Second Street, San Jose, CA 95113 (US).
- (72) Inventor: **HILLER, Dean**; 3599 Frost Road, Shrub Oak, NY 10588 (US).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: **METHOD AND APPARATUS FOR PROCESSING INTERNET SITE NAMES THROUGH REGULAR EXPRESSION COMPARISON**



(57) Abstract: A method and apparatus is presented where similar Internet site names (Fig. 2) are handled in an improved manner at a Domain Name Server (17). In one example, a first site name from a computer system (11) coupled to the Internet (17) is compared to a variety of site names through a regular expression comparison (29). Thus, a site that has a number of similar names (and potentially new site names that will have a similar format) will be better able to handle attempted accesses to site names that should logically go to its site.

WO 01/42943 A1



Published:

— With international search report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

**Method and Apparatus for Processing Internet Site Names
Through Regular Expression Comparison**

Field of the Invention

The present invention pertains to network communications. More particularly, the present
5 invention pertains to the processing of network addresses and site names (such as those found in the Internet) using regular expressions.

Background of the Invention

A network system includes a number of computer systems and local networks connected in
such a manner to allow communication between them. The Internet, which includes the so-called
10 World Wide Web, is an example of a global network system. A computer system coupled to the Internet (e.g., via a dial-up connection to an Internet Service Provider (ISP)) may communicate with any of the computer systems coupled in the Internet system. In other words, data, such as a request for information, may be directed to a particular computer system on the Internet.

Because of the number of computer systems in the network, a system is present to properly
15 route data communicated by one computer system to another. This is assisted in the Internet system with a plurality of Domain Name Servers. These servers communicate directly with a subset of the servers that are coupled in the Internet system. When a first computer seeks to send data intended for a second computer, a request is made to a first Domain Name Server to find the address of the second computer system on the Internet. As is known in the Internet art, the second computer
20 system is identified by a unique site name such as something having the form "abrainc.com." The first Domain Name Server that receives the request from the first computer system checks its stored site name information to see if there is an exact match for the "abrainc.com" character stream. If there is not a match, the first DNS sends a responsive message back to the first computer system. If there is a match, the first DNS forwards the request to another DNS where a similar comparison
25 between the "abrainc.com" string and those strings stored at the DNS. If there is a match at this DNS, associated with the match is a detailed Internet address (typically a string of numbers and periods) that is sent back to the first computer system. After the Internet address is supplied to the first computer system, the first computer system then sends its data, request, etc. to the second computer. Through this system, data, requests, etc. are forwarded between nodes in the Internet
30 until it reaches its destination.

In actuality, each unique Internet site name has associated with it a numerical Internet address. The DNS system is a hierarchical system. There are a small number of root DNSs which

are responsible for tracking the top level domains such as .com, .net, .edu, etc. These servers are instrumental in identifying valid site names. The system for locating properly addressed computer systems is a very robust system. The goal of the system is to forward requests to server systems that are more knowledgeable about the appropriate location of the DNS. The information that is stored
5 in the DNSs is repeatedly updated so that future requests to a particular site name can be properly handled.

One problem associated with this system is in how it handles similar site names (i.e., site names that have a similar number of characters or have character strings that are in common). In the art, similar site names are handled in exactly the same manner as for every other site name. In other
10 words, though similar site names may be directed to the same computer system, the DNS system must be updated for each and every site name to the computer system. This leads to inefficiency, in that it takes some time for the DNS system to be updated with the correct information so that newly created site names can be quickly located for communication.

Summary of the Invention

This and other problems are addressed by the method and apparatus of the present invention. According to an embodiment of the present invention a Domain Name Server is modified so as to use a regular expression in the comparison between the requested site name and the series of similar
15 site names (whether actual or potential) associated with a computer system.

Brief Description of the Drawings

Fig. 1 is a block diagram of a computer network system operated according to an
20 embodiment of the present invention.

Fig. 2 is a flow diagram of a method for operating a computer system of Fig. 1 according to an embodiment of the present invention.

Detailed Description

Referring to Fig. 1, a computer network that may be operated according to an embodiment of the present invention is shown. In this system, a first computer system 11 seeks to communicate with a second computer system 19. In doing so, it seeks to send some sort of information which will be referred to herein as "data" to a first site name which may be associated with computer
25 system 19. In this example, the first site name is similar to a plurality of other site names and is prone for searching using a regular expression for comparison purposes as described in further detail below. For the purposes of this embodiment, the site name is as follows: 4085551234.abrainc.com
30 (e.g., a U.S. telephone number with an "abrainc" sub-domain and a ".com" domain).

The first site name is communicated, for example, by the first computer system 11 to an Internet Service Provider (ISP) 13. The ISP 13 then transfers the site name to a Network Server 15 (e.g., over Internet 14). As known in the art, this Network Server may be a root Domain Name Server which is responsible for knowing valid site names in the ".com" realm. Once Network
5 Server 15 identifies a valid ".com" site name (i.e., the "abrainc.com" portion of the site name), the site name is eventually forwarded to the Internet and preferably is forwarded to Domain Name Server 17 to determine if 4085551234.abrainc.com is a valid site name (e.g., via a network node 18).

Referring to Fig. 2, a method of the present invention is shown in flow diagram form. In
10 block 21, a normal name lookup is performed where the first site name is compared to the site names that are stored at the DNS 17. At decision block 23 it is determined whether a match has been found. If one has, control passes to block 24 where an appropriate response is sent by the DNS 17 to first computer system 11. If no match is found, control passes to decision block 25 where it is determined whether regular expression comparisons are enabled at the DNS 17. If they are not, then
15 control passes to block 26 where an "unfound" message or the like is sent back to first computer system 11 (Fig. 1).

If regular expression comparisons are enabled, control passes to block 27 where a first regular expression is retrieved (e.g., from memory at DNS 17). According to an embodiment of the present invention, the first regular expression can have a form such as $^{\wedge}\text{d}\{10\}\$.X.Y$ where
20 $^{\wedge}\text{d}\{10\}\$$ represents a string of ten numbers, X represents a sub-level domain ("abrainc" in this example), and Y represents a top-level domain ("com" in this example). An alternative regular expression could be $^{\wedge}[0-9]+\$.X.Y$ where $^{\wedge}[0-9]+\$$ represents a string of numbers (each between 0 and 9). In block 29 a comparison is made between the first site name and the regular expression. Such a comparison can be done, for example using a comparison procedure such as that which is
25 used in the UNIX operating system. In this embodiment, the regular expressions set forth above are in a UNIX format. One skilled in the art based on the above teaching will appreciate that other formats may be used to represent a regular expression. In decision block 31, it is determined whether there is a match. If there is, then control passes to block 33 where a good response is sent to first computer system 11 with an appropriate Internet Protocol (IP) address. If there is not a
30 match, then control passes to decision block 35 to determine whether other regular expressions exist. If there are no additional regular expressions present at DNS 17, then control passes to block 26 where an "unfound" message or the like is sent to computer system 11. If there are additional

regular expressions, then control passes to block 37 to retrieve the next regular expression and pass control to block 29 to perform the next comparison.

Although several embodiments are specifically illustrated and described herein, it will be appreciated that modifications and variations of the present invention are covered by the above teachings and within the purview of the appended claims without departing from the spirit and intended scope of the invention. For example, Network Server 15 as a Domain Name Server may be modified to include a regular expression comparison as described above with reference to Fig. 1. In addition, other Domain Name Servers may be modified to include the regular expression method described above.

For example, the root Domain Name Server for a particular country may be used to perform the regular expression comparison. Accordingly, assuming that Z is a valid country code (i.e., a geographically-oriented top-level domain), all site names including that root are directed to the Domain Name Server for that country. Thus, modifying the regular expression above to `^d{10}$.Z` would result in the data being sent to the root Domain Name Server and that server determining that the site name is valid at that location. One skilled in the art will appreciate that other regular expressions may be used. For example regular expressions may be used which recognizes a string of alpha-numeric characters in the first site name.

What is claimed is:

1. A method of processing an Internet site name comprising:
performing a regular expression comparison between a first Internet site name and a
character pattern at a Domain Name Server.
- 5 2. The method of claim 1 further comprising:
transmitting said first Internet site name from a first computer system to said Domain
Name Server over the Internet.
3. The method of claim 2 further comprising:
transmitting a responsive message to said first computer system if a match is found in
10 said regular expression comparison.
4. The method of claim 1 wherein said regular expression uses a UNIX regular
expression format.
5. The method of claim 4 wherein said regular expression has a format $\backslash d\{10\}$.X.Y$
where X represents a sub-level domain and Y represents a top-level domain.
- 15 6. The method of claim 4 wherein said regular expression has a format $\wedge[0-9]+\$.X.Y$
where X represents a sub-level domain and Y represents a top-level domain.
7. The method of claim 4 wherein said regular expression has a format $\backslash d\{10\}$.Z$
where Z represents a geographically oriented top-level domain.
8. The method of claim 4 wherein said regular expression has a format $\wedge[0-9]+\$.Z$
20 where Z represents a geographically oriented top-level domain.
9. An apparatus for processing an Internet site name comprising:
a Domain Name Server adapted to perform a regular expression comparison between
a first site name and a character pattern.
10. A set of instructions residing in a storage medium, said set of instructions capable of
25 being executed by a processor to implement a method of processing an Internet site
name, the method comprising:
performing a regular expression comparison between a first Internet site name and a
character pattern at a Domain Name Server.
11. The set of instructions of claim 10, the method further comprising:
30 transmitting said first Internet site name from a first computer system to said Domain
Name Server over the Internet.
12. The set of instructions of claim 11, the method further comprising:

transmitting a responsive message to said first computer system if a match is found in said regular expression comparison.

13. The set of instructions of claim 10 wherein said regular expression uses a UNIX regular expression format.
14. The set of instructions of claim 13 wherein said regular expression has a format $\backslash d\{10\}$.X.Y where X represents a sub-level domain and Y represents a top-level domain.$
15. The set of instructions of claim 13 wherein said regular expression has a format $^{[0-9]}+$.X.Y where X represents a sub-level domain and Y represents a top-level domain.$
16. The set of instructions of claim 13 wherein said regular expression has a format $\backslash d\{10\}$.Z where Z represents a geographically oriented top-level domain.$
17. The set of instructions of claim 13 wherein said regular expression has a format $^{[0-9]}+$.Z where Z represents a geographically oriented top-level domain.$

1/2

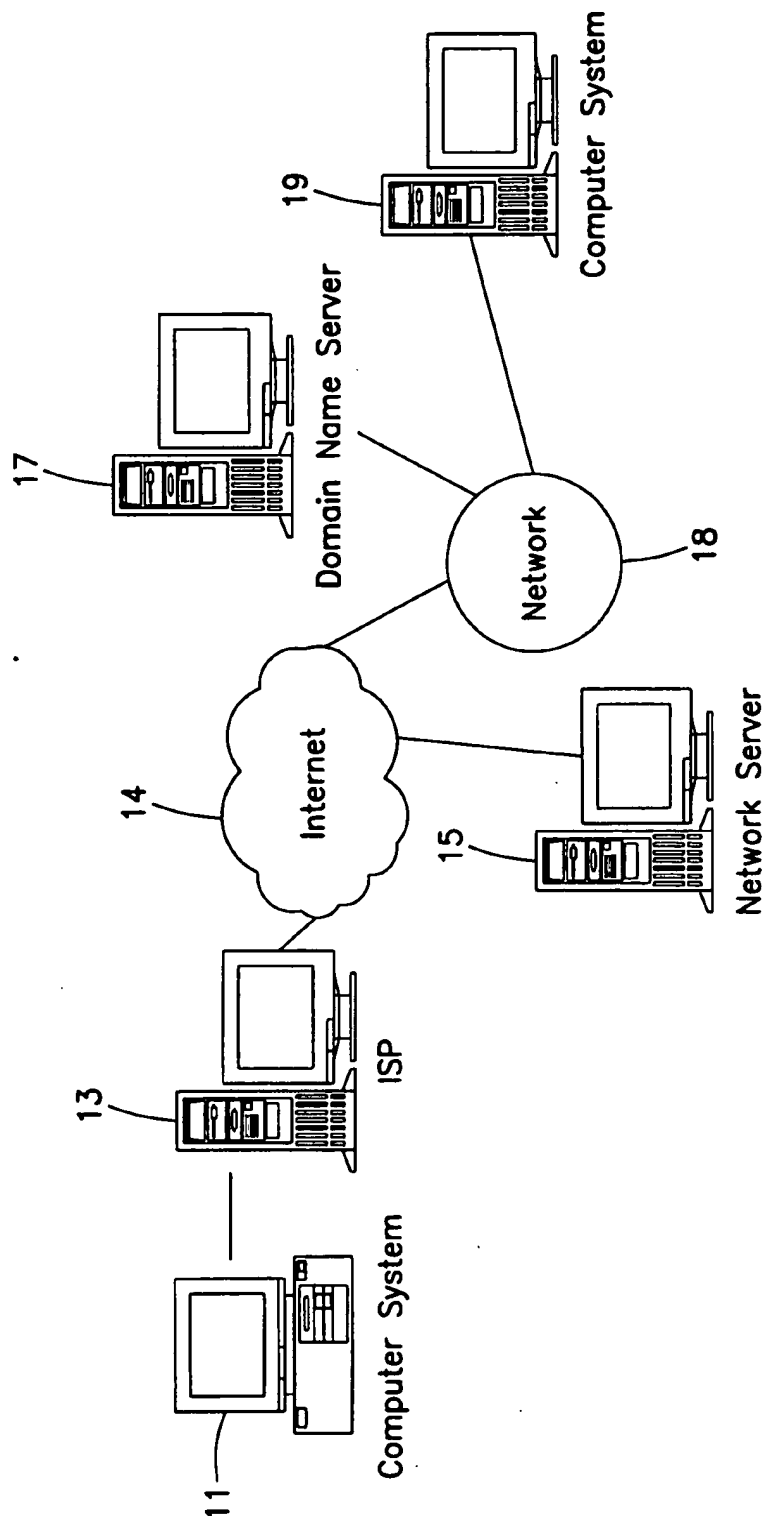
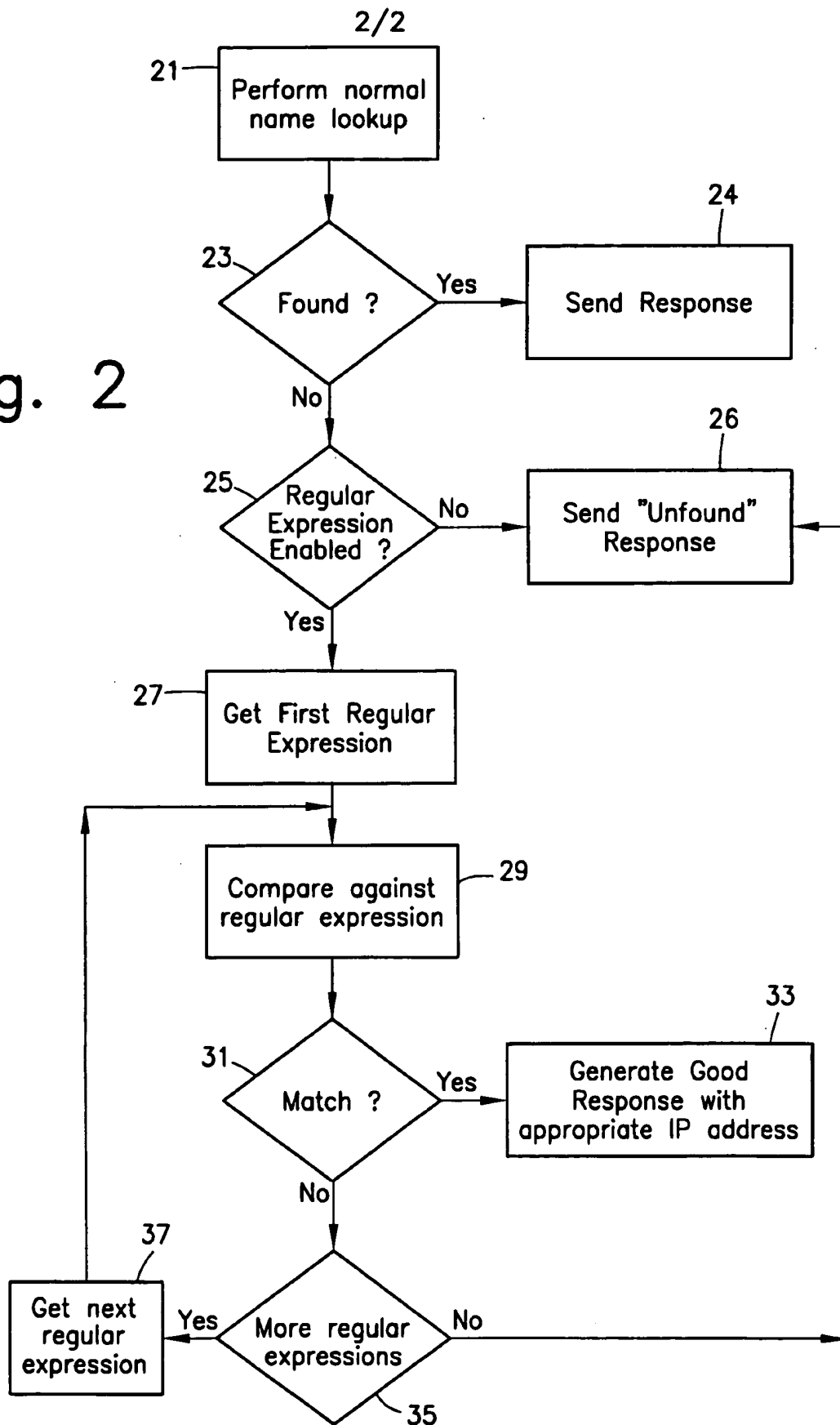


Fig. 1

Fig. 2



SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/09095

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06P 15/16

US CL : 709/244, 245

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 709/244, 245

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Microsoft Press COMPUTER DICTIONARY

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EAST, yahoo, altavista, excite

search terms: domain name system, dns, regular expression, UNIX, pattern match*

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	US 6,038,601 A (LAMBERT et al.) 14 March 2000, abstract, col. 20, lines 43-60	1-17
A, P	US 6,035,326 A (MILES et al.) 07 March 2000, abstract, col. 3, line 60 - col. 4, line 4, lines 23-45	1-17
A	author unknown, DNS, http://www.webopedia.com , May 1998, 3 pages	1-17
A	ORNAN ERES et al., DNS - The Domain Name System, Tel-Aviv University, http://www.rad.com/networks/1998/dns/welcome.htm , September 1998, 15 pages	1-17

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*g* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

21 JUNE 2000

Date of mailing of the international search report

11 JUL 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

PATRICE WINDER

Telephone No. (703) 305-3938

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/09095

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KRISTEN WINDBIGLER, Exploring the Domain Name Space, http://www.hotwired.lycos.com/webmonkey/webmonkey/geektalk/97/03/index4a.html , 24 January 1997, 4 pages, especially page 3	1-17
A	author unknown, Cisco DistributedDirector 2500 Series Installation and Configuration Guide, Chapter 1: Overview of the Cisco DistributedDirector 2500 Series, December 1997, 1-1 thur. 1-12	1-17
Y	author unknown, Cisco DistributedDirector 2500 Series Installation and Configuration Guide, Chapter 6: Configuring DNS Caching name server mode, December 1997, pages 6-1 thur. 6-18, especially 6-6	1-17
Y	author unknown, "regular expression", Webcom Online Users Guide, http://www.webcom.com , January 1999, 3 pages, especially pages 2-3	4-8, 13-17

Form PCT/ISA/210 (continuation of second sheet) (July 1998) ★